



Intel[®] Omni-Path Software

Release Notes for V10.9.3.1

Rev. 3.0

September 2019



You may not use or facilitate the use of this document in connection with any infringement or other legal analysis concerning Intel products described herein. You agree to grant Intel a non-exclusive, royalty-free license to any patent claim thereafter drafted which includes subject matter disclosed herein.

No license (express or implied, by estoppel or otherwise) to any intellectual property rights is granted by this document.

All information provided here is subject to change without notice. Contact your Intel representative to obtain the latest Intel product specifications and roadmaps.

The products described may contain design defects or errors known as errata which may cause the product to deviate from published specifications. Current characterized errata are available on request.

Intel technologies' features and benefits depend on system configuration and may require enabled hardware, software or service activation. Performance varies depending on system configuration. No computer system can be absolutely secure. Check with your system manufacturer or retailer or learn more at [intel.com](https://www.intel.com).

Intel, the Intel logo, Intel Xeon Phi, and Xeon are trademarks of Intel Corporation in the U.S. and/or other countries.

*Other names and brands may be claimed as the property of others.

Copyright © 2019, Intel Corporation. All rights reserved.



Contents

- 1.0 Overview of the Release..... 5**
 - 1.1 *Important Information*..... 5
 - 1.2 Audience..... 5
 - 1.3 Software License Agreement..... 5
 - 1.4 If You Need Help..... 5
 - 1.5 Enhancements and New Features in this Release..... 5
 - 1.6 Supported Features..... 6
 - 1.7 Deprecated and Removed Features..... 7
 - 1.8 Release Packages..... 7
 - 1.9 Release Compatibility..... 7
 - 1.10 Operating Systems..... 7
 - 1.11 Kernel Modules for hfi1 Driver..... 9
 - 1.12 Parallel File Systems..... 9
 - 1.13 MPI Libraries..... 10
 - 1.13.1 Supported MPI Libraries..... 10
 - 1.13.2 Compiler Versions and Distributions..... 10
 - 1.14 Intel Hardware..... 11
 - 1.15 Switch Firmware..... 11
 - 1.16 Document Versions..... 12
 - 1.17 Installation Requirements..... 12
 - 1.17.1 Software and Firmware Requirements..... 13
 - 1.17.2 OS RPMs Installation Prerequisites..... 13
 - 1.17.3 Installation Instructions..... 17
 - 1.17.4 Installation Path Changes in Release 10.4 (and later)..... 17
 - 1.18 Product Constraints..... 18
 - 1.19 Product Limitations..... 18
 - 1.20 Accelerated RDMA Information..... 19
- 2.0 Issues..... 20**
 - 2.1 Resolved Issues..... 20
 - 2.1.1 Issues Resolved in this Release..... 20
 - 2.1.2 Issues Resolved in Prior Releases..... 22
 - 2.2 Open Issues..... 23
 - 2.2.1 Third Party Open Issues..... 27
- 3.0 Related Information..... 28**
 - 3.1 Intel® Omni-Path Documentation Library..... 28
 - 3.1.1 How to Search the Intel® Omni-Path Documentation Set..... 30



Tables

1	Supported Operating Systems.....	8
2	Supported OS by CPU.....	8
3	Kernel Modules for hfi1 Driver.....	9
4	Supported MPI Libraries.....	10
5	Compiler Versions and Distributions.....	10
6	Supported Hardware.....	11
7	Supported Document Versions.....	12
8	RHEL* 7.5 Distribution RPMs.....	13
9	RHEL* 7.6 Distribution RPMs.....	14
10	RHEL* 8 Distribution RPMs.....	14
11	SLES* 12.3 Distribution RPMs.....	15
12	SLES* 12.4 Distribution RPMs.....	15
13	SLES* 15 Distribution RPMs.....	16
14	SLES* 15.1 Distribution RPMs.....	16
15	Issues Resolved in this Release.....	20
16	Issues Resolved in Prior Releases.....	22
17	Open Issues.....	23
18	Third Party Open Issues.....	27



1.0 Overview of the Release

This document provides a brief overview of the changes introduced into the Intel® Omni-Path Software by this release. References to more detailed information are provided where necessary. The information contained in this document is intended as supplemental information only; it should be used in conjunction with the documentation provided for each component.

These Release Notes list the features supported in this software release, open issues, and issues that were resolved during release development.

1.1 Important Information

This release resolves the following issue (see [Issues Resolved in this Release](#) on page 20):

- [STL-58309](#)

Please refer to open issues [STL-57800](#) and [STL-58303](#) in [Table 17](#) on page 23 for important information.

1.2 Audience

The information provided in this document is intended for installers, software support engineers, service personnel, and system administrators.

1.3 Software License Agreement

This software is provided under license agreements and may contain third-party software under separate third-party licensing. Please refer to the license files provided with the software for specific details.

1.4 If You Need Help

Technical support for Intel® Omni-Path products is available 24 hours a day, 365 days a year. Please contact Intel Customer Support or visit <http://www.intel.com/omnipath/support> for additional detail.

1.5 Enhancements and New Features in this Release

The following enhancements and features are new in this release:

- New operating systems supported. See [Table 1](#) on page 8 for details.
- Support for yum/zypper integration.
- Accelerated IPoFabric (AIP) performance improvement for IP routing and general use of IPoIB over Intel® Omni-Path Fabrics.



- Support for AIP provides improved performance for running IPoIB in datagram mode on Intel® Omni-Path fabrics. **Note that AIP is on by default and requires additional tuning.** See the *Intel® Omni-Path IP and LNet Router Design Guide* and *Intel® Omni-Path Fabric Performance Tuning User Guide* for more details.
- AIP is currently supported only in RHEL* 7.5, RHEL* 7.6, RHEL* 8, SLES* 12 SP4, SLES* 15 and SLES* 15 SP1.
- Support for libfabric OFI Version 1.6.2.
- Support for Sandia* OpenSHMEM 1.4.2.
- Support for NVIDIA* CUDA* 9.2 and 10.
- Support for Intel® Parallel Studio XE 2019.

1.6 Supported Features

- The list of supported operating systems is in [Table 1](#) on page 8.
- The list of supported hardware is in [Table 6](#) on page 11.
- Product Constraints described in [Product Constraints](#) on page 18.
- Active Optical Cables (AOC). For details, see the Cable Matrix at: <http://www.intel.com/content/www/us/en/high-performance-computing-fabrics/omni-path-cables.html>
- Support for the Enhanced Hypercube Routing Engine is outside the scope of Intel® OPA support. However, Intel partners may offer such support as part of their solutions. In addition there is an open source community who may be able to answer specific questions and provide guidance with respect to the Enhanced Hypercube Routing Engine.
- Support for Accelerated RDMA, also called Token ID (TID) RDMA, which is a Verbs protocol extension. See [Accelerated RDMA Information](#) on page 19 for details.
- Support for active optical cables (AOC) on server platforms using integrated HFI for OPA (commonly known as "-F").
- Support for GPUDirect* RDMA, which is a technology that enables a direct path for data exchange between a graphics processing unit (GPU) and a third-party peer device using standard features of PCI Express.
- Legacy BIOS Boot Mode Enhancements to support boot over fabric, custom board descriptions, and pre-boot platform configuration data for AOC support.
- Multi-endpoint functionality. See the *Intel® Performance Scaled Messaging 2 (PSM2) Programmer's Guide* for details.
- SNMP MIB support. See the *Intel® Omni-Path Fabric Switches Release Notes* for details.
- UEFI, TMM, and Firmware Tools delivered as standalone rpms.
- Support for Intel® C/C++ Compiler 18.0
- Intel® OPA support for cgroups.
- Support for multiple virtual fabric security.
- SE Linux PSM2 support with 13 labels.



1.7 Depreciated and Removed Features

- Support for the use of ESM and HSM as redundant pairs (i.e., an ESM as a backup to an HSM and vice versa) in a fabric has been deprecated in this release.
- In an upcoming release, native verbs support in Open MPI (openib BTL) will no longer be maintained. It will be removed from the Open MPI build in IFS.

1.8 Release Packages

There are two Intel® Omni-Path Fabric Software packages:

- Basic for compute nodes
- IFS for management nodes

The Basic package includes:

- Software that installs the following packages to the distribution OpenFabrics Alliance* (OFA):
 - hfi1-firmware, libpsm2 (for all RHEL*) and libpsm2-2 (for all SLES* 12 SP3 and newer), hfi1-diagtools-sw
 - Open MPI and MVAPICH2. See [MPI Libraries](#) on page 10 for details.
 - mpitests
 - mpi-selector
 - Sandia* OpenSHMEM
 - Open Fabrics Interface (OFI) libfabric

The IFS package includes the Basic package plus:

- Fabric Manager, which allows comprehensive control of administrative functions using a mature Subnet Manager. Fabric Manager simplifies subnet, fabric, and individual component management, easing the deployment and optimization of large fabrics.
- FastFabric Toolset, which enables rapid, error-free installation and configuration of Intel® OPA host software and management software tools, as well as simplified installation, configuration, validation, and optimization of HPC fabrics.

1.9 Release Compatibility

This release is backward compatible with the most recent minor release version. For example, Release 10.N is backward compatible with Release 10.N-1 and Release 10.N-1.x.

1.10 Operating Systems

This release of the Intel® Omni-Path Software supports the operating systems listed in the following table.

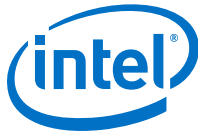


Table 1. Supported Operating Systems

Operating System	Update/SP	Kernel Version
Red Hat* Enterprise Linux* (RHEL*) 7.5 X86_64	Update 5	3.10.0-862.el7.x86_64
Red Hat* Enterprise Linux* (RHEL*) 7.6 X86_64	Update 6	3.10.0-957.el7.x86_64
Red Hat* Enterprise Linux* (RHEL*) 8 X86_64		4.18.0-80.el8.x86_64
CentOS*-7 (1804) X86_64 (corresponds to RHEL* 7.5)	(1804)	3.10.0-862.el7.x86_64
CentOS*-7 (1810) X86_64 (corresponds to RHEL* 7.6)	(1810)	3.10.0-957.el7.x86_64
Scientific Linux* 7.5 X86_64	Update 5	3.10.0-862.el7.x86_64
Scientific Linux* 7.6 X86_64	Update 6	3.10.0-957.el7.x86_64
SUSE* Linux* Enterprise Server (SLES*) 12.3 X86_64	Service Pack 3	4.4.114-94.14_default
SUSE* Linux* Enterprise Server (SLES*) 12.4 X86_64	Service Pack 4	4.12.14-94.41_default
SUSE* Linux* Enterprise Server (SLES*) 15 X86_64		4.12.14-23_default
SUSE* Linux* Enterprise Server (SLES*) 15.1 X86_64	Service Pack 1	4.12.14-195_default

Table 2. Supported OS by CPU

CPU Type	RHEL*			SLES*			
	7.5	7.6	8	12.3	12.4	15	15.1
Intel® Xeon Phi™ x200 Product Family (Knights Landing CPU-based servers)	•	•	•	•	•	•	•
Intel® Xeon® Scalable Processors (Skylake CPU-based servers)	•	•	•	•	•	•	•
2nd Generation Intel® Xeon® Scalable Processors (codename Cascade Lake)	•	•	•	•	•	•	•

CPU Type	CentOS*		Scientific Linux*	
	7 (1804)	7 (1810)	7.5	7.6
Intel® Xeon Phi™ x200 Product Family (Knights Landing CPU-based servers)	•	•	•	•
Intel® Xeon® Scalable Processors (Skylake CPU-based servers)	•	•	•	•
2nd Generation Intel® Xeon® Scalable Processors (codename Cascade Lake)	•	•	•	•

NOTE

PSM2 GPUDirect* RDMA with CUDA is supported as shown below:

- CUDA Toolkit 9.2 is supported on SLES* 12 SP3.
- CUDA Toolkit 10.0 is supported on RHEL* 7.5, RHEL* 7.6, and SLES* 15.
- Intel does not officially support CUDA with SLES* 12 SP4, SLES* 15 SP1, and RHEL* 8 at this time.



1.11 Kernel Modules for hfi1 Driver

This release of the Intel® Omni-Path Software contains the hfi1 driver kernel modules listed in the following table.

NOTE

The version number is shown in bold.

Table 3. Kernel Modules for hfi1 Driver

OS	Non-CUDA/CUDA	RPM Version
RHEL* 7.5	Non-CUDA	kmod-ifs-kernel-updates-3.10.0_862.el7.x86_64- 1792 .x86_64.rpm
	CUDA	kmod-ifs-kernel-updates-3.10.0_862.el7.x86_64- 1792cuda .x86_64.rpm
RHEL* 7.6	Non-CUDA	kmod-ifs-kernel-updates-3.10.0_957.el7.x86_64- 1792 .x86_64.rpm
	CUDA	kmod-ifs-kernel-updates-3.10.0_957.el7.x86_64- 1792cuda .x86_64.rpm
RHEL* 8	Non-CUDA	kmod-ifs-kernel-updates-4.18.0_80.el8.x86_64- 1792 .x86_64.rpm
SLES* 12.3	Non-CUDA	ifs-kernel-updates-kmp-default-4.4.114_94.14_default_k4.4.114_94.14- 1792 .x86_64.rpm
	CUDA	ifs-kernel-updates-kmp-default-4.4.114_94.14_default_k4.4.114_94.14- 1792cuda .x86_64.rpm
SLES* 12.4	Non-CUDA	ifs-kernel-updates-kmp-default-4.12.14_94.41_default_k4.12.14_94.41- 1792 .x86_64.rpm
SLES* 15	Non-CUDA	ifs-kernel-updates-kmp-default-4.12.14_23_default_k4.12.14_23- 1792 .x86_64.rpm
	CUDA	ifs-kernel-updates-kmp-default-4.12.14_23_default_k4.12.14_23- 1792cuda .x86_64.rpm
SLES* 15.1	Non-CUDA	ifs-kernel-updates-kmp-default-4.12.14_195_default_k4.12.14_195- 1792 .x86_64.rpm

1.12 Parallel File Systems

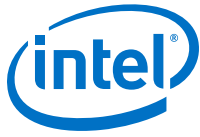
This section reports the parallel file systems that have been tested with this release of the Intel® Omni-Path Software. Note that not all combinations of OS, file system, and Intel® OPA software are tested.

The following parallel file systems have been tested:

- Lustre* Long Term Support (LTS) release:
 - Version 2.10.4 on RHEL* 7.5
 - Version 2.10.6 on SLES* 12 SP3 (via LNET Self Test)
- IBM* Spectrum Scale¹:
 - Version 5.0.2.3 on RHEL* 7.5, RHEL* 7.6, SLES* 12 SP3

Refer to the *Intel® Omni-Path Fabric Performance Tuning User Guide* for details on optimizing parallel file system performance with Intel® Omni-Path Software.

¹ Formerly known as General Parallel File System (GPFS)



1.13 MPI Libraries

In the last release, MPI builds have been reduced to the libraries as listed in [Table 4](#).

NOTE

The sources for Open MPI and MVAPICH2 are included in the software package. FastFabric provides a simple tool to build these MPIs using the Intel® Compiler (not included in the software package). Refer to *Intel® Omni-Path Fabric Suite FastFabric User Guide*, [Rebuilding MPI Library and Tools](#).

1.13.1 Supported MPI Libraries

The table below lists the different MPI libraries supported by Intel® Omni-Path Fabric Software with the corresponding version, fabric support, and compiler used. Note that the second column indicates whether the MPI library is included in the Intel® Omni-Path Software package or not.

Table 4. Supported MPI Libraries

MPI Implementation	Included in Basic Package?	Runs Over	Compiled With
Open MPI 2.1.2	Yes	PSM2, OFI	GCC
Open MPI 2.1.2-cuda	Yes	PSM2	GCC
MVAPICH2-2.3B	Yes	PSM2	GCC
Intel® MPI Library 2018 Update 2	No	PSM2 (via TMI), OFI	N/A
Intel® MPI Library 2019, Initial Release	No	PSM2 (via OFI)	N/A
IBM* Spectrum* MPI version 10.1.1.0	No	PSM2	N/A

1.13.2 Compiler Versions and Distributions

The MPI libraries listed in the preceding section that are included in the release and built with PSM2 support were built with the following compiler versions:

Table 5. Compiler Versions and Distributions

Compiler	OS Distribution	Compiler Version
(GNU) gcc	RHEL* 7.5	gcc version 4.8.5 20150623 (Red Hat* 4.8.5-28) (GCC)
(GNU) gcc	RHEL* 7.6	gcc version 4.8.5 20150623 (Red Hat* 4.8.5-36) (GCC)
(GNU) gcc	RHEL* 8	gcc version 8.2.1 20180905 (Red Hat* 8.2.1-3) (GCC)
(GNU) gcc	SLES* 12 SP3	gcc version 4.8.5 (SUSE* Linux*)
(GNU) gcc	SLES* 12 SP4	gcc version 4.8.5 (SUSE* Linux)
(GNU) gcc	SLES* 15	gcc version 7.3.1 20180323 (SUSE* Linux*)
(GNU) gcc	SLES* 15 SP1	gcc version 7.4.1 20190424 [gcc-7-branch revision 270538] (SUSE* Linux*)

**NOTE**

Refer to the *Intel® Omni-Path Fabric Host Software User Guide* for setup information when using Open MPI with the SLURM PMI launcher and PSM2.

1.14 Intel Hardware

The following table lists the Intel hardware supported in this release. The table does not include OEM-specific hardware, such as custom adapters and switches.

NOTE

The Intel® PSM2 implementation has a limit of four (4) HFIs.

Table 6. Supported Hardware

Hardware	Description
Intel® Xeon® Processor E5-2600 v3 product family	Haswell CPU-based servers
Intel® Xeon® Processor E5-2600 v4 product family	Broadwell CPU-based servers
Intel® Xeon® Scalable Processors	Skylake CPU-based servers
2nd Generation Intel® Xeon® Scalable Processors	Cascade Lake CPU-based servers
Intel® Xeon Phi™ x200 Product Family	Knights Landing CPU-based servers
Intel® Omni-Path Host Fabric Interface 100HFA016 (x16)	Single Port Host Fabric Interface (HFI)
Intel® Omni-Path Host Fabric Interface 100HFA018 (x8)	Single Port Host Fabric Interface (HFI)
Intel® Omni-Path Switch 100SWE48Q	Managed 48-port Edge Switch
Intel® Omni-Path Switch 100SWE48U	Externally-managed 48-port Edge Switch
Intel® Omni-Path Switch 100SWE48UFH	Externally-managed 48-port Edge Switch, hot-swap power and fans
Intel® Omni-Path Switch 100SWE48QFH	Managed 48-port Edge Switch, hot-swap power and fans
Intel® Omni-Path Switch 100SWE24Q	Managed 24-port Edge Switch
Intel® Omni-Path Switch 100SWE24U	Externally-managed 24-port Edge Switch
Intel® Omni-Path Director Class Switch 100SWD24	Director Class Switch 100 Series, up to 768 ports
Intel® Omni-Path Director Class Switch 100SWD06	Director Class Switch 100 Series, up to 192 ports

1.15 Switch Firmware

The following firmware is supported for Intel® Omni-Path switches:

- Intel® Omni-Path Switch Firmware 10.8.x revision (managed and externally-managed switches)
- Intel® Omni-Path Switch Firmware 10.7.x revision (managed and externally-managed switches)

Refer to the *Intel® Omni-Path Fabric Switches Release Notes* for more information.



1.16 Document Versions

The following table lists the end user document versions supported by this release.

Table 7. Supported Document Versions

Title	Doc. Number	Revision
Intel® Omni-Path Fabric Quick Start Guide	J57479	6.0
Intel® Omni-Path Fabric Setup Guide	J27600	10.0
Intel® Omni-Path Fabric Switches Hardware Installation Guide	H76456	10.0
Intel® Omni-Path Host Fabric Interface Installation Guide	H76466	5.0
Intel® Omni-Path Fabric Software Installation Guide	H76467	14.0
Intel® Omni-Path Fabric Switches GUI User Guide	H76457	10.0
Intel® Omni-Path Fabric Switches Command Line Interface Reference Guide	H76458	10.0
Intel® Omni-Path Fabric Suite FastFabric User Guide	H76469	14.0
Intel® Omni-Path Fabric Suite Fabric Manager User Guide	H76468	13.0
Intel® Omni-Path Fabric Suite Fabric Manager GUI User Guide	H76471	13.0
Intel® Omni-Path Fabric Host Software User Guide	H76470	14.0
Intel® Performance Scaled Messaging 2 (PSM2) Programmer's Guide	H76473	12.0
Intel® Omni-Path Fabric Performance Tuning User Guide	H93143	16.0
Intel® Omni-Path IP and LNet Router Design Guide (Old title: Intel® Omni-Path IP and Storage Router Design Guide)	H99668	8.0
Building Containers for Intel® Omni-Path Fabrics using Docker* and Singularity* Application Note	J57474	7.0
Intel® Omni-Path Management API Programmer's Guide	J68876	6.0
Configuring Non-Volatile Memory Express* (NVMe*) over Fabrics on Intel® Omni-Path Architecture Application Note	J78967	1.0
Intel® Omni-Path Fabric Software Release Notes	K59645	1.0
Intel® Omni-Path Fabric Manager GUI Release Notes	K61052	1.0
Intel® Omni-Path Fabric Switches Release Notes (includes managed and externally- managed switches)	K38337	1.0
Intel® Omni-Path Fabric Unified Extensible Firmware Interface (UEFI) Release Notes	K50782	1.0
Intel® Omni-Path Fabric Thermal Management Microchip (TMM) Release Notes	K38341	1.0
Intel® Omni-Path Fabric Firmware Tools Release Notes	K38342	1.0

Related Links

[Intel Omni-Path Documentation Library](#) on page 28

1.17 Installation Requirements



1.17.1 Software and Firmware Requirements

Table 1 on page 8 lists the operating systems supported by this release. For the required packages, refer to [OS RPMs Installation Prerequisites](#).

1.17.2 OS RPMs Installation Prerequisites

Ensure that the following requirements are met prior to installing the software.

- In addition to normal OS installation options, OS RPMs must be installed before you can install the Intel® Omni-Path software.
- Refer to the applicable section below to verify that all required RPMs are installed for the specific version of your OS distribution.
- Depending on the packages you choose, there may be additional prerequisites. For additional information, refer to the Release Notes for your specific release and installation type.

Red Hat* Enterprise Linux* OS RPMs

The tables below list the RPMs for each supported RHEL* release.

NOTES

- Some RHEL* rpms are available in a Server-Optional repository. Please install the RHEL* Server-Optional version from Red Hat* which contains additional, required development packages.
- Rdma-ndd is part of rdma-core.

Table 8. RHEL* 7.5 Distribution RPMs

IB/OPA Centric	System Centric		Other/Generally Installed	Build Requirements
ibacm	atlas	libstdc++-devel	bash	expat-devel
infinipath-psm	bc	ncurses-libs	irqbalance	libnl3-devel
libibcm	coreutils	numactl-libs	kernel	libpfm
libibmad	createrepo	openssl	kmod	libuuid-devel
libibumad	expat	openssl-devel	libgcc	numactl-devel
libibverbs	expect	openssl-libs	perl	opensm-libs
libnl3	gcc-gfortran	pciutils	perl-Getopt-Long	openssl-devel (1.0.1 or higher)
librdmacm	glibc	redhat-rpm-config	perl-Socket	valgrind-devel
opensm-libs	kernel-devel	rpm-build	pkgconfig	
perftest	libgfortran	sysfsutils	python	
qperf	libgomp	tcl	systemd	
rdma-core	libquadmath	tcsh		
rdma-core-devel	libstdc++	zlib		



Table 9. RHEL* 7.6 Distribution RPMs

IB/OPA Centric	System Centric		Other/Generally Installed	Build Requirements
ibacm	atlas	libstdc++-devel	bash	expat-devel
infinipath-psm	bc	ncurses-libs	irqbalance	libnl3-devel
libibumad	coreutils	numactl-libs	kernel	libpfm
libibverbs	createrepo	openssl	kmod	libuuid-devel
libnl3	expat	openssl-devel	libgcc	numactl-devel
librdmacm	expect	openssl-libs	perl	opensm-libs
opensm-libs	gcc-gfortran	pciutils	perl-Getopt-Long	openssl-devel (1.0.1 or higher)
perftest	glibc	redhat-rpm-config	perl-PathTools	valgrind-devel
qperf	kernel-devel	rpm-build	perl-Socket	
rdma-core	libgfortran	sysfsutils	pkgconfig	
rdma-core-devel	libgomp	tcl	python	
	libquadmath	tcsh	systemd	
	libstdc++	zlib	systemd-libs	

Table 10. RHEL* 8 Distribution RPMs

IB/OPA Centric	System Centric		Other/Generally Installed	Build Requirements
ibacm	atlas	libstdc++-devel	bash	autoconf
libibumad	bc	ncurses-libs	irqbalance	automake
libibverbs	coreutils	numactl-libs	kernel	expat-devel
librdmacm	createrepo	openssl	kernel-modules-extra	gcc-c++.x86_64
opensm-libs	expat	openssl-devel	kmod	kernel-abi-whitelists
perftest	expect	openssl-libs	libgcc	kernel-rpm-macros
qperf	gcc-gfortran	pciutils	perl	libnl3-devel
rdma-core	glibc	redhat-rpm-config	perl-Getopt-Long	libpfm
rdma-core-devel	kernel-devel	rpm-build	perl-Socket	libtool
	libatomic	sysfsutils	pkgconf	libuuid-devel
	libgfortran	tcl	python2	numactl-devel
	libgomp	tcsh	systemd	opensm-libs
	libquadmath	zlib		openssl-devel (1.1.1 or higher)
	libstdc++			

SUSE* Linux* Enterprise Server OS RPMs

The tables below list the RPMs for each supported SLES* release.



NOTE

Some SLES* rpms are available in SLES* Software Development Kit (SDK). Please install the SLES* SDK iso from SUSE* which contains additional, required development packages.

Table 11. SLES* 12.3 Distribution RPMs

IB/OPA Centric	System Centric		Other/Generally Installed	Build Requirements
ibacm	bc	libnuma1	bash	libexpat-devel
libibcm1	createrepo	libopenssl1_0_0	coreutils	libnuma-devel
libibmad5	expect	libopenssl-devel	glibc	libopenssl-devel (1.0.1 or higher)
libibumad3	gcc-fortran	libquadmath0	grep	libuuid-devel
libibverbs1	kernel-devel	libudev-devel	irqbalance	opensm-libs3
libnl3-200	kernel-syms	libz1	kmod	valgrind-devel
libpsm_infinipath1	libexpat1	openssl	libedit0	
librdmacm1	libgfortran3	rpm-build	libgcc_s1	
mpi-selector	libgomp1	tcl	libstdc++6	
opensm-devel	libncurses5	tclsh	libudev1	
opensm-libs3			perl	
perftest			perl-base	
qperf			pkg-config	
rdma-core			python-base	
rdma-core-devel			systemd	
rdma-ndd			udev	

Table 12. SLES* 12.4 Distribution RPMs

IB/OPA Centric	System Centric		Other/Generally Installed	Build Requirements
ibacm	bc	libnuma1	bash	libexpat-devel
libibcm1	createrepo	libopenssl1_0_0	glibc	libnuma-devel
libibmad5	expect	libopenssl-devel	grep	libopenssl-devel (1.0.1 or higher)
libibumad3	gcc-fortran	libquadmath0	irqbalance	libuuid-devel
libibverbs1	kernel-devel	libudev-devel	kmod	opensm-libs3
libpsm_infinipath1	kernel-syms	libz1	libedit0	valgrind-devel
librdmacm1	libexpat1	openssl	libgcc_s1	
mpi-selector	libgfortran3	rpm-build	libstdc++6	
opensm-devel	libgomp1	tcl	perl	

continued...



IB/OPA Centric	System Centric		Other/Generally Installed	Build Requirements
opensm-libs3	libncurses5	tcsch	perl-base	
perftest			pkg-config	
qperf			python-base	
rdma-core			systemd	
rdma-core-devel			udev	
rdma-ndd				

Table 13. SLES* 15 Distribution RPMs

IB/OPA Centric	System Centric		Other/Generally Installed	Build Requirements
ibacm	bc	libnuma1	bash	kernel-devel
libibmad5	coreutils	libopenssl1_1	glibc	libexpat-devel
libibumad3	createrepo_c	libopenssl-devel	grep	libnuma-devel
libibverbs1	expect	libosmcomp3	irqbalance	libopenssl-devel (1.0.1 or higher)
libnl3-200	gcc-fortran	libquadmath0	kmod	libuuid-devel
libpsm_infinipath1	kernel-devel	libudev-devel	libedit0	opensm-libs3
libquadmath0	kernel-syms	libz1	libgcc_s1	valgrind-devel
librdmacm1	libexpat1	openssl	libncurses5	
libverbs1	libgfortran4	rpm-build	libstdc++6	
mpi-selector	libgomp1	tcl	perl	
opensm-devel	libncurses6	tcsch	perl-base	
opensm-libs3			pkg-config	
perftest			python-base	
qperf			systemd	
rdma-core			udev	
rdma-core-devel				
rdma-ndd				

Table 14. SLES* 15.1 Distribution RPMs

IB/OPA Centric	System Centric		Other/Generally Installed	Build Requirements
ibacm	bc	libnuma1	bash	kernel-devel
libibmad5	coreutils	libopenssl1_1	glibc	libexpat-devel
libibumad3	createrepo_c	libopenssl-devel	grep	libnuma-devel
libibverbs1	expect	libosmcomp3	irqbalance	libopenssl-devel (1.0.1 or higher)
libnl3-200	gcc-fortran	libquadmath0	kmod	libuuid-devel

continued...



IB/OPA Centric	System Centric		Other/Generally Installed	Build Requirements
libpsm_infinipath1	kernel-devel	libudev-devel	libedit0	opensm-libs3
libquadmath0	kernel-syms	libz1	libgcc_s1	valgrind-devel
librdmacm1	libexpat1	openssl	libncurses5	
libverbs1	libgfortran4	rpm-build	libstdc++6	
mpi-selector	libgomp1	tcl	perl	
opensm-devel	libncurses6	tcsh	perl-base	
opensm-libs3			pkg-config	
perftest			python-base	
qperf			systemd	
rdma-core			udev	
rdma-core-devel				
rdma-ndd				

Example

The example below shows the install of a group of OS RPMs on a RHEL* OS server. Not all required OS RPMs are included in this example and some OS RPMs in this example might not be needed.

```
# yum install libibmad libibverbs librdmacm libibcm qperf perftest
rdma infinipath-psm expat libstdc++-devel gcc-gfortran atlas tcl expect
tcsh sysfsutils pciutils bc libibumad libibumad-devel libibumad
libibumad-devel libibverbs-devel libibmad-devel librdmacm-devel ibacm-devel
openssl-devel libuuid-devel expat-devel infinipath-psm-devel valgrind-devel
libgnome libibverbs* opensm-libs libhfil ncurses-devel hwloc hwloc-gui
```

Intel recommends that you build your own list of OS RPMs for installation.

1.17.3 Installation Instructions

There are two Intel® Omni-Path Fabric Software packages:

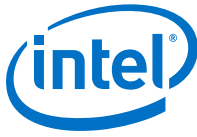
- IntelOPA-IFS.<distro>-x86_64.<version>.tgz for the management node.
- IntelOPA-Basic.<distro>-x86_64.<version>.tgz for compute nodes.

The packages in the tgz file are RPMs. Installing individual RPMs is not supported in this release.

Refer to the *Intel® Omni-Path Fabric Software Installation Guide* for related software requirements and complete installation procedures. Refer to the *Intel® Omni-Path Fabric Switches Hardware Installation Guide* for related firmware requirements.

1.17.4 Installation Path Changes in Release 10.4 (and later)

If you are upgrading an Intel® Omni-Path Fabric Software installation (Release 10.3 or earlier), Intel recommends that you perform the following steps before upgrading, due to changes in installation paths for RPMs and configuration files.



Run `./INSTALL -u` to uninstall existing packages.

Run `./INSTALL -a` to complete the installation.

Pre-existing configuration files are automatically saved by the RPM as `.rpmsave` files. (RPM will notify you about these files during removal.) If you want to keep these configuration files, you should move them to their new locations. A mapping of old configuration file locations to new locations is shown in the following table.

Old Location (Release 10.3 and earlier)	New Location (Release 10.4 and later)
<code>/etc/sysconfig/opafm.xml</code>	<code>/etc/opa-fm/opafm.xml</code>
<code>/etc/sysconfig/allhosts</code>	<code>/etc/opa/allhosts</code>
<code>/etc/sysconfig/chassis</code>	<code>/etc/opa/chassis</code>
<code>/etc/sysconfig/esm_chassis</code>	<code>/etc/opa/esm_chassis</code>
<code>/etc/sysconfig/hosts</code>	<code>/etc/opa/hosts</code>
<code>/etc/sysconfig/opafastfabric.conf</code>	<code>/etc/opa/opafastfabric.conf</code>
<code>/etc/sysconfig/opaff.xml</code>	<code>/etc/opa/opaff.xml</code> Changed to <code>/etc/opa/opamgt_tls.xml</code> in 10.7.
<code>/etc/sysconfig/opamon.conf</code>	<code>/etc/opa/opamon.conf</code>
<code>/etc/sysconfig/ports</code>	<code>/etc/opa/ports</code>
<code>/etc/sysconfig/switches</code>	<code>/etc/opa/switches</code>

1.18 Product Constraints

- Power class 2 AOC are supported. You must use 10.5 (or newer) host software and 1.5 (or newer) UEFI for proper operation. Integrated HFI (-F) requires a specific BIOS level to support power class 2 AOC; contact your BIOS vendor for more information.
- Special considerations are required if using *Intel® Parallel Studio 2019, Update 2*.
 - The `libpsm2` library provided in the Intel® OPA IFS package has been compiled with *Intel® Parallel Studio 2019, initial release*. Note that it does not compile correctly with the Intel® Parallel Studio 2019, Update 2.
Please use the *Intel® Parallel Studio 2019 - Initial release* compiler if you need to recompile the PSM2 library.
 - When linking OpenMPI against the IFS-provided version of `libfabric` and Intel® MPI Library 2019, Update 2 is installed, the following environment variable setting must be made to ensure proper linking with the IFS `libfabric`:

```
source <IntelMPI installdir>/intel64/bin/mpivars.sh -  
ofi_internal=0
```

1.19 Product Limitations

This release has the following product limitations:

- The embedded version of the Fabric Manager supports a maximum of 100 HFI ports involving less than 20 switch ASICs. Calculate the number of switch ASICs in your fabric as follows:



- One ASIC per Intel® Omni-Path Edge Switch 100 Series
- Two ASICs per Intel® Omni-Path Director Class Switch 100 Series Leaf module
- Two ASICs per Intel® Omni-Path Director Class Switch 100 Series Spine module
- Performance Administration (PA) Failover should **not** be enabled with FMs running on differing software versions.
To disable PA failover, edit the `/etc/sysconfig/opafm.xml` file and in the `<Pm>` section, change `<ImageUpdateInterval>` to 0.
- Enabling UEFI Optimized Boot on some platforms can prevent the HFI UEFI driver from loading during boot. To prevent this, do not enable UEFI Optimized Boot.

1.20 Accelerated RDMA Information

Accelerated RDMA is a Verbs protocol extension to improve the performance of RDMA write and RDMA read operations on Intel® Omni-Path hardware.

This extension improves the efficiency of large message transfers to provide performance benefits for storage protocols and other Verbs-based protocols. The benefits include increased achievable bandwidth with reduced CPU utilization. The Token ID (TID) RDMA protocol accelerates the OpenFabrics Alliance* (OFA) Verbs API with no changes required to API consumers. The acceleration technique is performed by the host driver and the application running over the OFA Verbs API does not need to make any code changes.

Accelerated RDMA is off by default.

To enable it, add `cap_mask=0x4c09a01cbba` to the `/etc/modprobe.d/hfi1.conf` file. Instructions on how to do this are in the *Intel® Omni-Path Fabric Performance Tuning User Guide*, Setting HFI1 Driver Parameters section.

NOTE

It is not necessary to enable Accelerated RDMA on all nodes in a cluster. The performance effects of enabling Accelerated RDMA on a subset of nodes has not been characterized. It is highly recommended that you test to see if enabling this feature works best for your applications.



2.0 Issues

This section lists the resolved and open issues in the Intel® Omni-Path Software.

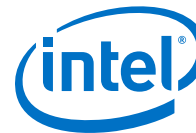
2.1 Resolved Issues

2.1.1 Issues Resolved in this Release

The following table lists issues that are resolved in this release.

Table 15. Issues Resolved in this Release

ID	Description	Resolved in Release
STL-58309	It is possible for a memory region not to be freed due to the lack of rcu synchronization for the memory region's refcount. This has been observed when executing the <code>mmvdisk</code> while running GPFS version RAID version 5.0.2-3.0.1.	10.9.3.1
STL-46789	SLES 12.2 shipped with an old version <code>ibacm</code> that does not include the unit file. Since IFS INSTALL relies on <code>systemd</code> to turn on/off services, enabling <code>ibacm</code> with IFS INSTALL fails. <i>Note:</i> SLES 12.2 is no longer supported.	10.9.3
STL-56887	When installing OPA with CUDA, <code>openmpi_gcc_cuda_hfi</code> is required. When creating a repository for OPA without CUDA, <code>opacreaterepo</code> script recommends <code>openmpi_gcc_cuda_hfi</code> by mistake.	10.9.3
STL-47578	Due to a SLES* 15 kernel setting, <code>hfi1_eprom</code> cannot work while the HFI driver is loaded. The tool and driver are mutually exclusive.	10.9.2
STL-55620	When Accelerated IPoFabric (AIP) is enabled, less cores are available for HFI driver <code>sdma</code> interrupts. The exact number depends on the number of CPU cores in the system. This could negatively impact send performance.	10.9.2
STL-56600	Using <code>./INSTALL -n</code> to do a fresh installation will not set the <code>irqbalance</code> policy. This can cause a performance issue.	10.9.2
STL-47546	When an ISL goes down in the middle of an FM sweep (due to a disruption in the fabric such as a reboot), the SA copy of topology becomes invalid when the Fatree routing algorithm is used. SA queries that use this topology (e.g., path record query) fail. <i>Note:</i> A path record query failure can be seen in FM log as "INVALID TOPOLOGY" messages. The issue will resolve after the FM's next successful sweep.	10.9.1
STL-46211 STL-47907	For SLES OSes, when IPoIB is enabled and working, the IFS INSTALL erroneously displays IPoIB as "disabled". When the IFS is upgraded, this display error causes the INSTALL to disable IPoIB in order to keep the same configuration.	10.9.1
STL-48065	In some boot scenarios the BIOS can order the device paths such that it passes the device handle of a device implementing the <code>EFI_NETWORK_INTERFACE_IDENTIFIER_PROTOCOL</code> , which is not a child of the HFI device path. In these cases, the HFI UEFI driver can attempt to obtain access to internal structures based on the physical memory location of the controller name passed. This can result in an unexpected behavior in the BIOS. This is addressed by explicitly confirming that the controller handle passed by the BIOS matches the HFI device controller for child device handles.	10.9.0.1
<i>continued...</i>		



ID	Description	Resolved in Release
132207	Kernel crash caused by the ib_srpt module.	10.9.0
139743 143031 143115	Under a very heavy load through the iPoIB interface, the kernel warning NETDEV WATCHDOG: ib0 (hfil): transmit queue 0 timed out, followed by the messages queue stopped 1, tx_head xxx, tx_tail xxx and transmit timeout: latency xxxx msecs may be seen. As of V10.9.0, this issue is no longer valid.	10.9.0
143296	When irqbalance uses the argument --hintpolicy=exact, it applies the policy of setting the hardware interrupts to CPU core mappings according to device drivers preferences. For the HFI1 driver, it is strongly recommended to preserve interrupt locality for low latency and high bandwidth by having a dedicated CPU core per interrupt.	10.9.0
141740	Third-Party Issue: Due to the race condition during boot up ipoib driver can miss PKey change event that will leave ipoib interface in disconnected state.	10.9.0
144996	Running workloads with more than 78 ranks with the Open MPI OFI MTL over OFI Verbs;OFI_RXM provider may result in a hang with message sizes larger than 65 KB.	10.9.0
145415	ESM hangs after the first sweep in a configuration that includes all <VirtualFabric> sections that are enabled and with QOS =1, where <BaseSL> is defined in all with values other than 0. Symptoms of this hang are: <ul style="list-style-type: none"> • SA or PA queries to ESM time out (e.g., opareport fails with FTIMEOUT) • SM does not resweep after SweepInterval, multicast membership change, or fabric change • smControl shutdown is issued but ESM never reaches stopped state 	10.9.0
145474	OFI Verbs mpi_stress may cause verbs/MSG provider completion queue overrun that results in dropped completions. They show up as sequence errors in the test.	10.9.0
145585	SLES* 15 ./INSTALL script does not properly run dracut -f.	10.9.0
145623	For systems running on SLES* 15, there is a known issue with irqbalance.	10.9.0
145855	If the Admin VF is not running on VL0, the HSM may get into a state where it is unable to talk to the fabric. The sweep will log the following errors: <pre> opamgt ERROR: [<pid>] omgt_send_mad2: send failed; Invalid argument, agent id 2 MClass 0x81 method 0x1 attrId 0x11 attrM 0x0 WARN [topology]: SM: sm_send_stl_request_impl: Error Sending to Path:[1] Lid:[0xffffffff] [Can't find node in topology!]. AID:[NODEINFO] TID:[0x0000000000000031] Status:[OK (0x00000000)] WARN [topology]: SM: topology_main: TT: too many errors during sweep, will re-sweep in a few seconds rc: 108: unrecoverable error </pre>	10.9.0
146271	When using libpsm2 with PSM2_CUDA=1, workloads may assert() during MPI_Recv or MPI_collective operations where the source buffer is a GPU buffer and located on the same node, i.e., intranode transfer. The destination buffer locality has no influence. This assert() will only affect transfers that are not aligned to the start of the source buffer.	10.9.0
146456	In a fabric with only one Edge switch using the fat tree routing algorithm, a port can get stuck in the Init (LinkUp) state after the port is bounced.	10.9.0
STL-46606 STL-47956 STL-48661	Bouncing a link or rebooting a device under certain fabric conditions may cause a switch in the fabric to be removed from the Fabric Manager's internal view of the topology leading to fabric disruptions and instability.	10.9.0
STL-46790	In cases where GSI services are active and the FM is receiving capability change traps (common after node reboots), FM responsiveness may be impacted. This could result in data traffic disruption or unexpected FM failovers. GSI traffic would include the PM, SA, and DBSync (FM failover).	10.9.0



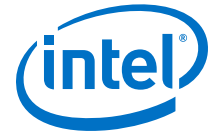
2.1.2 Issues Resolved in Prior Releases

The following table lists issues that were resolved in prior releases.

Table 16. Issues Resolved in Prior Releases

ID	Description	Resolved in Release
STL-46790	In cases where GSI services are active and the FM is receiving capability change traps (common after node reboots), FM responsiveness may be impacted. This could result in data traffic disruption or unexpected FM failovers. GSI traffic would include the PM, SA, and DBSync (FM failover).	10.8.0.2
STL-46606 STL-47956 STL-48661	Bouncing a link or rebooting a device under certain fabric conditions may cause a switch in the fabric to be removed from the Fabric Manager's internal view of the topology leading to fabric disruptions and instability.	10.8.0.1
STL-47546	When an ISL goes down in the middle of an FM sweep (due to a disruption in the fabric such as a reboot), the SA copy of topology becomes invalid when the Fattree routing algorithm is used. SA queries that use this topology (e.g., path record query) fail. <i>Note:</i> A path record query failure can be seen in FM log as "INVALID TOPOLOGY" messages. The issue will resolve after the FM's next successful sweep.	10.8.0.1
135981 143631	In RHEL 7.3, irqbalance spams messages every ten seconds with snippet below. The HFI driver sets affinity and the Omni Fabric install sets <code>/etc/sysconfig/irqbalance</code> to use exact. <code>IRQBALANCE_ARGS= --hintpolicy=exact</code> Irq balancer scans the <code>isolcpus</code> and <code>nohz_full</code> kernel masks and adds the corresponding CPUs to the <code>banned_cpus</code> mask. This works fine for valid masks, but not for the default, empty masks. In this case, when they read from the <code>sysfs</code> they return empty strings, <code>"\n"</code> or <code>"0x0, \n"</code> <i>Note:</i> RHEL* 7.3 is no longer supported.	10.8
136432	Certain perftest tools such as <code>ib_write_bw</code> do not work on RHEL* 7.3 when using the RDMA CM with UD QPs. <i>Note:</i> RHEL* 7.3 is no longer supported.	10.8
140797 141558 141852 142476	When installing an IFS tar package on a server where the OS-resident OPA software has been installed, the installation may fail to properly update all the RPMs. This is due to the use of Epoch tagging on the OS-resident OPA software which makes the OS-resident RPMs appear to be newer than RPMs in the IFS tar package.	10.8
141263	The GUPS application that is included as part of Sandia Open Shmem apps may fail with more than 1% of error when executed with more that four PEs. <i>Note:</i> GUPS does not use atomic XOR operations given they are not supported in the OpenShmem specification, version 1.3. This is addressed in version 1.4.	10.8
141636	During the Intel OPA installation, the ramdisk may be rebuilt multiple times; however, only the last build is necessary.	10.8
141586	Using SLES* 12 SP2 and SP3, <code>INSTALL -a</code> fails due to dependency issues. The following information to resolve this issue has been added to the <i>Intel® Omni-Path Fabric Software Installation Guide</i> , OS RPMs Installation Prerequisites section: "Prior to installing Intel OPA, install the SLES SDK iso from SUSE which contains additional, required development packages."	10.8
141793	Use of a static buffer could produce an incorrect device name (<code>hfi1_x</code>) in <code>dmesg</code> logging.	10.8
142330	MPI applications that leverage the PSM2 library's access to the HFI ASICs Memory Mapped IO and that access the MMIO directly (not via PSM2) can potentially cause an "unsupported opcode" error which some servers handle as a critical error. <i>Note:</i> A side effect of this issue is that the <code>hfi1</code> driver automatically masks the AER Unsupported Request error for the OPA1 HFI device.	10.8

continued...



ID	Description	Resolved in Release
143311	During UEFI pre-boot with the connected switch running 10.7.0.0.134, the OPA link may not complete initialization if the link is bounced or restarted. This behavior is limited to the pre-boot period. There is no exposure once Linux boot has completed.	10.8
143449	PM will scroll LQI=0 and Integrity Exceeded Threshold logs when an additional VF with QoS enabled and a device group that is not "All". <i>Note:</i> This issue does not occur when running against the default opafm.xml configuration file.	10.8
143915	When specific messages that are of non-double word length (not 4 B) and less than 1 MTU in size (10,240 B in most configurations) are sent, they may never complete, causing an application to wait indefinitely. <i>Note:</i> This issue occurs on machines with more than 28 CPU cores (not including hyper-threads) or more than 28 HFI contexts enabled per Intel® Omni-Path HFI.	10.8
144165	Nodes unable to ping on IPoIB. <i>Note:</i> This issue occurs when a host port disappears and reappears from the FM's topology (usually due discovery timeout or major fabric disruption), while the port remains ACTIVE the entire time. This results in the host port not being a member of the IP multicast groups. The primary symptom is the inability to resolve IP addresses via ARP.	10.8
144795	ESM assigns invalid MLIDs. IPoIB and other applications that use multicast will not function correctly with invalid MLIDs.	10.8
145399 146088	Using the <code>./INSTALL -i</code> command results in errors due to missing dependencies.	10.8
145917	As required by the System Management Bus (SMBus) specification, when the Bus is idle for less than 50 μ s, the HFI card pulls the Bus into low voltage. However, the Bus does not remain in low voltage for a long enough interval.	10.8

2.2 Open Issues

The following table lists the open issues for this release.

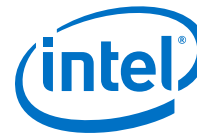
Table 17. Open Issues

ID	Description	Workaround
129563 (STL-47095)	Memory allocation errors with MVAPICH2-2.1/Verbs.	<i>Note:</i> To avoid this issue, use MPIs over PSM. If you are using MPIs over verbs, the following workaround is required: <ul style="list-style-type: none"> When running MVAPICH2 jobs with a large number of ranks (for example, > 36 ranks but \leq 72 ranks), you must set the following parameters in <code>/etc/security/limits.conf</code>: <ul style="list-style-type: none"> hard memlock unlimited soft memlock unlimited Also, you must increase the <code>lkey_table_size:LKEY</code> table size in bits (2^n, where $1 \leq n \leq 23$) from its default of 16 to 17. For instructions on setting module parameters, refer to the <i>Intel® Omni-Path Fabric Performance Tuning User Guide</i>, HFI1 Driver Module Parameters chapter.
134819	In KNL-F EFI shell, the command <code>ifconfig -l</code> does not correctly display the IP address after being assigned via DHCP.	Launch a newer version of the EFI shell from the embedded shell.
135028	NVMe over Fabric Protocol is only supported on Intel® OPA with Linux* kernel 4.5 and later versions.	To use NVMe functionality on Intel® OPA, you must patch the kernel.
<i>continued...</i>		



ID	Description	Workaround
135830 (STL-46193)	On Intel® Xeon Phi™ systems, failure observed during software upgrade when rebuilding the boot image. Error message contains: Rebuilding boot image with "/usr/bin/dracut -f"	Install parallel gzip (pigz) and make it the default compression tool prior to upgrading the software.
136822	The Intel UEFI driver contained in the server BIOS must be executed for proper support of Active Optical Cables (AOC) in an integrated HFI environment. Some BIOS do not execute the UEFI in Legacy BIOS Boot mode, and there are BIOS configuration settings that may prevent the UEFI from executing in any mode.	Avoid the use of Legacy BIOS boot mode if your platform does not execute the HFI driver in that mode. Avoid BIOS settings or other configuration settings that do not execute the HFI driver during boot.
139368	Some applications compiled with older compilers may use a personality bit that signifies that READ should imply EXECUTE permissions. To improve system security, the hfi1 driver does not allow execute permissions on PSM memory maps. Therefore, applications that use READ implies EXECUTE will fail to run.	As root, run the execstack tool to clear the executable bit on the binary: <code>execstack -c <binary></code> Alternatively, recompile the binary to not set this personality bit.
139613	The Subsystem Vendor and Subsystem Device ID in the PCI configuration space of Intel® Omni-Path discrete HFI cards may not indicate the correct OEM vendor and device. As a result, the <code>lspci</code> command may show incorrect Subsystem Vendor and Device ID information. This issue affects Intel server boards for Intel® Xeon® Processor v3 and v4 Product Family configured in Legacy OS boot mode.	Reconfigure the system from Legacy OS boot mode to UEFI boot mode.
139995	When installing Intel® Omni-Path Software on a clean system (with no prior versions installed), the following error message may be displayed: <code>cat: //etc/opa/version_delta: No such file or directory</code>	This message can be safely ignored. The installer is looking for an IFS version file before it has been created on the system. The installation is not impacted.
141005 (STL-47901)	In 10.8, the output of the module parameter <code>num_user_contexts</code> has been changed. The value of the module parameter <code>num_user_context</code> used at driver init time determines the number of receive contexts reserved for PSM. The default value of -1 caused the driver init sequence to determine how many CPUs are available, and assigned the context count to that value. <i>It would then update the module parameter <code>num_user_contexts</code> with that value.</i> This incorrect behavior has been fixed. <code>num_user_contexts</code> is no longer updated.	Do not use <code>num_user_contexts</code> to determine how many context can be used. To determine the number of available user contexts, refer to the sysfs file: <code>/sys/class/infiniband/hfi1_<port>/nctxts</code>
143174 (STL-47003)	Due to a SLES* 15 kernel setting, hfidiaqs cannot work while the HFI driver is loaded. The tool and driver are mutually exclusive.	Boot the kernel with <code>iomem=relaxed</code> . This turns off the resource exclusivity check.
145771	Due to changes in the SLES* 15 kernel, a user space application cannot access a hardware resource if that resource is being used by a kernel driver. The result is that the <code>hfi1_eprom</code> cannot access the EEPROM on an HFI when <code>hfi1</code> kernel driver is using the device.	Perform the following workaround: 1. Unload <code>hfi1</code> kernel driver: <code>rmmmod hfi1</code> 2. Run <code>hfi1_eprom</code> commands. 3. Reload <code>hfi1</code> kernel driver: <code>modprobe hfi1</code>
STL-43764	An FM running on SLES* 12.3 on certain older CPUs may experience instability due to a CPU hardware errata.	Follow instructions provided by SUSE to disable hardware lock elision at: https://www.suse.com/support/kb/doc/?id=7022289

continued...



ID	Description	Workaround
STL-47125	In rare cases, a program crash may be experienced during process cleanup when running with the OpenMPI PSM2 MTL (Matching Transport Layer).	Restart the workload.
STL-47571	<p>Since libfabric 1.6 (which is included in IFS 10.8), the psm2 provider maps OFI endpoints directly to HFI contexts instead of multiplexing multiple OFI endpoints to a single HFI context. This relies on the multi-EP feature of the PSM2 library and thus the provider automatically sets PSM2_MULTI_EP=1 if it has not been set. However, enabling the multi-EP feature also disables context sharing. As the result, applications may experience the following runtime error when trying to oversubscribe CPU cores (which is usually the same as available HFI contexts).</p> <pre data-bbox="386 684 776 758">hfi_userinit: assign_context command failed: Device or resource busy PSM2 can't open hfi unit: -1 (err=23)</pre> <p><i>Note:</i> Applications that don't use libfabric are not affected.</p>	<p>Set PSM2_MULTI_EP=0.</p> <p><i>Note:</i> This only works for applications that open only one OFI endpoint per process.</p>
STL-48480	When running workloads with Intel® MPI or OpenMPI using OFI RXM and Verbs providers, if more than 100 ranks are running, it is possible for the workload to hang during message passing.	<p>If Verbs messaging is not required, use one of the following methods to run with the PSM2 provider:</p> <ul data-bbox="914 894 1414 989" style="list-style-type: none"> • Intel® MPI: <code>-genv I_MPI_OFI_PROVIDER 'psm2'</code> • OpenMPI: <code>--mca mtl_ofi_provider_include 'psm2'</code>
STL-48921	<p>If you attempt to unload or restart the hfi1 driver while the ib_ipoib driver is loaded, it may fail with a message similar to this:</p> <pre data-bbox="386 1087 862 1108">modprobe: FATAL: Module hfi1 is in use</pre>	To avoid this issue, unload the ib_ipoib driver before unloading or restarting the hfi1 driver.
STL-49210	For RHEL* 7.5 and newer and SLES* 15, SLES* 12.4 and newer, when Accelerated IP (AIP) is disabled and IPoFabric is configured in datagram mode, no bulk traffic will occur if the configured ib0 MTU size is larger than what default (non-AIP) IPoFabric can support.	Enable AIP or reduce the locally configured ib0 MTU size to a size that default IPoFabric can support (e.g., 4092 bytes).
STL-49313	On OS versions (RHEL* 7.5 and newer and SLES* 15, SLES* 12.4 and newer) where it is available, configuring IPoIB module parameter <code>ipoib_enhanced=0</code> disables AIP on same node.	To run with AIP, ensure enhanced IPoIB is enabled (<code>ipoib_enhanced=1</code>).
STL-56236	When running workloads with OpenMPI using the OFI RXM and Verbs providers, a rare segmentation fault may be encountered during message processing with many ranks. This issue will result in the workload failing and will require the workload to be rerun.	<p>If Verbs messaging is not required, PSM2 messaging can be used to workaround this issue by running:</p> <pre data-bbox="914 1444 1373 1465">--mca mtl_ofi_provider_include 'psm2'</pre>
STL-56557	<p>The primary IPoIB network device associated with any RDMA device may fail to join certain multicast groups. This can prevent IPv6 neighbor discovery and possibly other network ULPS from working correctly. Note that the IPv4 broadcast group is not affected as the IPoIB network device handles joining that multicast group directly.</p> <p>This problem does not affect IPoIB child network devices.</p>	<p>Perform one of the following workarounds:</p> <ul data-bbox="914 1556 1390 1671" style="list-style-type: none"> • Do not use IPv6 over a parent IPoIB device. • Only load the IPoIB module after the HFI link is ACTIVE with a valid pkey assigned. • Bounce (unload, then load) the IPoIB module anytime a pkey change occurs.

continued...



ID	Description	Workaround
STL-57040	If Intel® MPI 2019, Update 2 and OpenMPI are installed on the same cluster, OpenMPI jobs could link to the Intel® MPI internal libfabric library after running Intel® MPI mpivars.sh.	<p>Before running an OpenMPI job, perform the following to reset the libfabric environment variables that had been previously set by Intel® MPI.</p> <pre>source <installdir>/intel64/bin/mpivars.sh - ofi_internal=0</pre>
STL-57127	Restarting the irqbalance service while the HFI1 driver is loaded may cause improper or inconsistent distribution of interrupts, which may result in low network transfer performance.	<p>Irqbalance must be stopped before starting the HFI1 module. Perform the following command sequence:</p> <pre>systemctl stop irqbalance modprobe -r ib_ipoib modprobe -r hfi1 modprobe hfi1 modprobe ib_ipoib systemctl start irqbalance</pre>
STL-57790	A known issue with OpenMPI 2.1.2 and one-sided operations (such as MPI_Accumulate) when running under RHEL 8 results in an MPI internal error.	Download and install OpenMPI 2.1.6 or newer on RHEL 8 fabrics.
STL-57800 STL-58303	INSTALL enables ARPTABLE_TUNING by default. With this flag, Intel® OPA modifies /etc/sysctl.conf for arp table size adjustment. If you modify this file after ARPTABLE_TUNING is enabled, your changes are not permanent. Rebooting the machine will restore the file to Intel® OPA's version.	<p>If you do not need ARPTABLE_TUNING, perform either of the following:</p> <ul style="list-style-type: none"> • If OPA is not installed, run the Install TUI and answer n (no) when asked whether to apply ARPTABLE_TUNING or not. • If OPA is already installed, run the following cmd to turn off ARPTABLE_TUNING: <pre>opaautostartconfig --disable ARPTABLE_TUNE</pre> <p>If you need ARPTABLE_TUNING and you need to edit the /etc/sysctl.conf file, perform the following:</p> <ol style="list-style-type: none"> 1. With OPA already installed, run the following cmd to stop ARPTABLE_TUNING: <pre>opa-arptbl-tuneup stop</pre> 2. Edit /etc/sysctl.conf with new settings. 3. Run the following cmd to start ARPTABLE_TUNING: <pre>opa-arptbl-tuneup start</pre> 4. Reboot the machine. <p>Your settings and OPA's arp table size adjustment will both appear in the sysctl.conf file.</p>
STL-57994	The OSU benchmarks built and packaged with the CUDA-enabled MPI packages included in the IFS software distribution are not built with CUDA support.	Download the OSU benchmarks from http://mvapich.cse.ohio-state.edu/benchmarks/ . Follow the instructions for enabling CUDA when building.
STL-58332	When upgrading from RHEL 7.5 to a newer version, if a value is set for ifs_sel_mode in hfi1.conf HFI driver will not load.	Remove the parameter and value.
STL-58338	Under RHEL 8, installing mpitests_openmpi_gcc_hfi with yum fails because no package provides libmpi.so.20() (64-bit). This shared library is shipped in openmpi_gcc_hfi; however, the openmpi_gcc_hfi rpm fails to claim that it provides libmpi.so.20.	<p>To install mpitests_openmpi_gcc_hfi, perform the following steps:</p> <ol style="list-style-type: none"> 1. Install openmpi_gcc_hfi. 2. Do a force installation of mpitests_openmpi_gcc_hfi without a dependency check. <p>For example:</p> <pre>rpm -ivh --nodeps mpitests_openmpi_gcc_hfi-4.0-931.x86_64.rpm</pre>



2.2.1 Third Party Open Issues

The following table lists the third party open issues for this release.

Table 18. Third Party Open Issues

ID	Description	Workaround
133604	Bonding driver shows incorrect hardware address of IPoIB interfaces.	This workaround applies to SLES* 12 SP3 (and earlier) distributions: Use the <code>opainfo</code> command to retrieve the PortGUID and <code>ip addr show ib0</code> to get the correct 20-byte hardware address of OPA network interface.
139924	For SLES*, the <code>ibacmp</code> provided in the OS distribution uses incorrect address information when joining multicast groups. This causes name resolution to fail. The <code>dsap</code> provided in the OS distribution works correctly.	The fix for this issue is available in the library <code>rdma-core-15-2.el7</code> . <ul style="list-style-type: none"> Do not use the OS distribution <code>ibacmp</code>. Install <code>rdma-core-15-2.el7</code> manually.
140310 (STL-46700)	On RHEL* 7.5, if an <code>ipofabric</code> interface name is non-standard, the <code>ifcfg</code> file is not read correctly. For example, if you are using the <code>ipofabric</code> interface name <code>opa_ib0</code> , and the <code>connected</code> mode and MTU size is specified in the <code>ifcfg-opa_ib0</code> file, when you bring up the interface with <code>ifup</code> , the settings do not take effect.	You must manually enter: <code>echo connected > /sys/class/net/opa_ib0/mode</code> for the settings to take effect.
141273 (STL-46935)	The in-distro version of <code>perftests</code> has bugs.	Use the upstream version of <code>perftest</code> from https://github.com/linux-rdma/perftest .



3.0 Related Information

3.1 Intel® Omni-Path Documentation Library

Intel® Omni-Path publications are available at the following URLs:

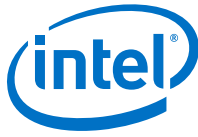
- Intel® Omni-Path Switches Installation, User, and Reference Guides
<http://www.intel.com/omnipath/SwitchPublications>
- Intel® Omni-Path Software Installation, User, and Reference Guides (includes HFI documents)
<http://www.intel.com/omnipath/FabricSoftwarePublications>
- Drivers and Software (including Release Notes)
<http://www.intel.com/omnipath/Downloads>

Use the tasks listed in this table to find the corresponding Intel® Omni-Path document.

Task	Document Title	Description
Key: Shading indicates the URL to use for accessing the particular document.		
• Intel® Omni-Path Switches Installation, User, and Reference Guides: http://www.intel.com/omnipath/SwitchPublications		
• Intel® Omni-Path Software Installation, User, and Reference Guides (includes HFI documents): http://www.intel.com/omnipath/FabricSoftwarePublications (no shading)		
• Drivers, Software, and Firmware (including Release Notes): http://www.intel.com/omnipath/Downloads		
Using the Intel® OPA documentation set	<i>Intel® Omni-Path Fabric Quick Start Guide</i>	A roadmap to Intel's comprehensive library of publications describing all aspects of the product family. This document outlines the most basic steps for getting your Intel® Omni-Path Architecture (Intel® OPA) cluster installed and operational.
Setting up an Intel® OPA cluster	<i>Intel® Omni-Path Fabric Setup Guide</i>	Provides a high level overview of the steps required to stage a customer-based installation of the Intel® Omni-Path Fabric. Procedures and key reference documents, such as Intel® Omni-Path user guides and installation guides, are provided to clarify the process. Additional commands and best known methods are defined to facilitate the installation process and troubleshooting.
Installing hardware	<i>Intel® Omni-Path Fabric Switches Hardware Installation Guide</i>	Describes the hardware installation and initial configuration tasks for the Intel® Omni-Path Switches 100 Series. This includes: Intel® Omni-Path Edge Switches 100 Series, 24 and 48-port configurable Edge switches, and Intel® Omni-Path Director Class Switches 100 Series.
	<i>Intel® Omni-Path Host Fabric Interface Installation Guide</i>	Contains instructions for installing the HFI in an Intel® OPA cluster.
continued...		



Task	Document Title	Description
Installing host software Installing HFI firmware Installing switch firmware (externally-managed switches)	<i>Intel® Omni-Path Fabric Software Installation Guide</i>	Describes using a Text-based User Interface (TUI) to guide you through the installation process. You have the option of using command line interface (CLI) commands to perform the installation or install using the Linux* distribution software.
Managing a switch using Chassis Viewer GUI Installing switch firmware (managed switches)	<i>Intel® Omni-Path Fabric Switches GUI User Guide</i>	Describes the graphical user interface (GUI) of the Intel® Omni-Path Fabric Chassis Viewer GUI. This document provides task-oriented procedures for configuring and managing the Intel® Omni-Path Switch family. Help: GUI embedded help files
Managing a switch using the CLI Installing switch firmware (managed switches)	<i>Intel® Omni-Path Fabric Switches Command Line Interface Reference Guide</i>	Describes the command line interface (CLI) task information for the Intel® Omni-Path Switch family. Help: -help for each CLI
Managing a fabric using FastFabric	<i>Intel® Omni-Path Fabric Suite FastFabric User Guide</i>	Provides instructions for using the set of fabric management tools designed to simplify and optimize common fabric management tasks. The management tools consist of Text-based User Interface (TUI) menus and command line interface (CLI) commands. Help: -help and man pages for each CLI. Also, all host CLI commands can be accessed as console help in the Fabric Manager GUI.
Managing a fabric using Fabric Manager	<i>Intel® Omni-Path Fabric Suite Fabric Manager User Guide</i>	The Fabric Manager uses a well defined management protocol to communicate with management agents in every Intel® Omni-Path Host Fabric Interface (HFI) and switch. Through these interfaces the Fabric Manager is able to discover, configure, and monitor the fabric.
	<i>Intel® Omni-Path Fabric Suite Fabric Manager GUI User Guide</i>	Provides an intuitive, scalable dashboard and set of analysis tools for graphically monitoring fabric status and configuration. This document is a user-friendly alternative to traditional command-line tools for day-to-day monitoring of fabric health. Help: Fabric Manager GUI embedded help files
Configuring and administering Intel® HFI and IPoIB driver Running MPI applications on Intel® OPA	<i>Intel® Omni-Path Fabric Host Software User Guide</i>	Describes how to set up and administer the Host Fabric Interface (HFI) after the software has been installed. The audience for this document includes cluster administrators and Message-Passing Interface (MPI) application programmers.
Writing and running middleware that uses Intel® OPA	<i>Intel® Performance Scaled Messaging 2 (PSM2) Programmer's Guide</i>	Provides a reference for programmers working with the Intel® PSM2 Application Programming Interface (API). The Performance Scaled Messaging 2 API (PSM2 API) is a low-level user-level communications interface.
Optimizing system performance	<i>Intel® Omni-Path Fabric Performance Tuning User Guide</i>	Describes BIOS settings and parameters that have been shown to ensure best performance, or make performance more consistent, on Intel® Omni-Path Architecture. If you are interested in benchmarking the performance of your system, these tips may help you obtain better performance.
Designing an IP or LNet router on Intel® OPA	<i>Intel® Omni-Path IP and LNet Router Design Guide</i> (Old title: <i>Intel® Omni-Path IP and Storage Router Design Guide</i>)	Describes how to install, configure, and administer an IPoIB router solution (Linux* IP or LNet) for inter-operating between Intel® Omni-Path and a legacy InfiniBand* fabric.
continued...		



Task	Document Title	Description
Building Containers for Intel® OPA fabrics	<i>Building Containers for Intel® Omni-Path Fabrics using Docker* and Singularity* Application Note</i>	Provides basic information for building and running Docker* and Singularity* containers on Linux*-based computer platforms that incorporate Intel® Omni-Path networking technology.
Writing management applications that interface with Intel® OPA	<i>Intel® Omni-Path Management API Programmer's Guide</i>	Contains a reference for programmers working with the Intel® Omni-Path Architecture Management (Intel OPAMGT) Application Programming Interface (API). The Intel OPAMGT API is a C-API permitting in-band and out-of-band queries of the FM's Subnet Administrator and Performance Administrator.
Using NVMe* over Fabrics on Intel® OPA	<i>Configuring Non-Volatile Memory Express* (NVMe*) over Fabrics on Intel® Omni-Path Architecture Application Note</i>	Describes how to implement a simple Intel® Omni-Path Architecture-based point-to-point configuration with one target and one host server.
Learning about new release features, open issues, and resolved issues for a particular release	<i>Intel® Omni-Path Fabric Software Release Notes</i>	
	<i>Intel® Omni-Path Fabric Manager GUI Release Notes</i>	
	<i>Intel® Omni-Path Fabric Switches Release Notes (includes managed and externally-managed switches)</i>	
	<i>Intel® Omni-Path Fabric Unified Extensible Firmware Interface (UEFI) Release Notes</i>	
	<i>Intel® Omni-Path Fabric Thermal Management Microchip (TMM) Release Notes</i>	
	<i>Intel® Omni-Path Fabric Firmware Tools Release Notes</i>	

3.1.1 How to Search the Intel® Omni-Path Documentation Set

Many PDF readers, such as Adobe* Reader and Foxit* Reader, allow you to search across multiple PDFs in a folder.

Follow these steps:

1. Download and unzip all the Intel® Omni-Path PDFs into a single folder.
2. Open your PDF reader and use **CTRL-SHIFT-F** to open the Advanced Search window.
3. Select **All PDF documents in...**
4. Select **Browse for Location** in the dropdown menu and navigate to the folder containing the PDFs.
5. Enter the string you are looking for and click **Search**.

Use advanced features to further refine your search criteria. Refer to your PDF reader Help for details.